REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	26.Oct.98	MAJOR REPORT

4. TITLE AND SUBTITLE
A COMPARISON OF CLASSIFICATION TECHNIQUES IN ENVIRONMENTAL APPLICATIONS OF REMOTE SENSING

5. FUNDING NUMBERS

6. AUTHOR(S)
2D LT BEAN MELANIE P

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF COLORADO AT COLORADO SPRINGS

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
98-032

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

A Comparison of Classification Techniques in Environmental Applications of Remote Sensing

Melanie Bean

B. S. E. University of Notre Dame, 1997

A creative investigation report submitted to the Faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirements for the degree of Master of Engineering in Space Operations, Spring 1998.

Abstract

This paper provides a comparison of remote sensing classification techniques as an extension of the environmental applications of remote sensing using Landsat data. It explores the history of remotes sensing, the principles of electromagnetic energy, and the general steps involved in remote sensing. Further, it describes the remote sensing features of Landsat -4, and -5. Finally, it applies classification techniques using Idrisi software.

Table of Contents

List of Figures and Tables

Introduction

This project began as an investigation into the use of Landsat remote sensing imagery for the monitoring of Fountain Creek Watershed in Colorado Springs, Colorado. The imagery was available, but not quite appropriate for this application. The focus then turned to Landsat's use in general environmental studies with an emphasis on comparing different types of classification available for use here at UCCS.

Simply gathering images through aerial photography or even space monitoring does not constitute remote sensing. These images are useless without the proper handling. A complete remote sensing application consists of three stages: 1. emission or reflection of energy from an object, 2. image capture, and 3. interpretation and analysis of image.

The first stage consists of the propagation of energy through the atmosphere, energy interactions with earth surface features, and retransmission of the energy through the atmosphere. Determining an active or passive remote sensing system depends on the radiation source. If the system relies on solar radiation, it is considered passive. If the system generates its own radiation such as radar or sonar, it is considered an active system. This paper is concerned with passive systems only.

The second stage can be accomplished through a variety of avenues. Aerial photography remains a popular means of obtaining high resolution images, but more and more satellite imagery systems are coming into play in both the commercial and government arenas. Two of the most prominent systems today are the European based SPOT satellite system, and NASA's Landsat program. Due to the availability of Landsat images to UCCS, only this image capturing system will be explored in this paper.

Interpretation and analysis of the image then finish the remote sensing process. This step integrates the image, error correction, and ground truth data. Software and data manipulation correct for any distortion in the image due to the atmosphere, motion of the satellite, flaws in the sensor, etc. They also serve to enhance certain features of the image for better interpretation. However, no interpretation can be correct without comparing it to ground truth, or actual physical data gathered at the site or sites like it. Combining the image, error correction, and ground truth, objects in the image can be classified. Through this process, every portion of the image receives a distinction as designated by the end user. This paper focuses on several forms of error correction, available types of ground truth, and a comparison of different classification techniques using Idrisi software.

1.0 Previous Environmental Studies

1.1 Hydrology Applications

Remote sensing can lend tremendous support to the monitoring of watersheds. It is particularly useful in gathering information about remote territories, such as water contained in snow packed or ice areas. Remote sensing also lends a hand in reducing the cost of repetitive and seasonal monitoring of lakes. For instance, monitoring the levels of the hundreds of lakes contained in even a small section of the upper mid-west can be time consuming and costly using conventional methods. However, the use of remote sensing as an assessment tool reduces both the time and expenses involved in this monitoring [1]. Due to scope of this paper, only water availability and flood monitoring will be discussed.

1.1.1 Water Availability

Water availability assessment through the use of Landsat imagery has had moderate success. A sharp contrast defines water masses from land in the infared region of the spectrum. Studies cited by Striffler report that Landsat imagery can be used to identify ninety-eight percent of all surface water in most areas. However, the bodies of water detected must be larger than at least one pixel on a Landsat image (30 m x 30 m). Ideally, the body would span several pixels for an accurate assessment of its cover. In fact, further studies cited by Striffler indicated that Landsat imagery was unsuitable for identifying bodies of water less than 80 meters wide. For this application, aerial photography, or systems with greater resolution are preferred.

1.1.2 Flood Monitoring

Landsat images have been used for the study of flood waters in many cases including the Mississippi River and the Indus River in Pakistan. In both these cases, the rivers studied can be easily identified on the Landsat imagery because of their immense proportions. An effective tool in the Indus River case was contrast stretching, which increased the difference between wet and dry areas, as well as delineating areas of leakage in dams and canals [1]. However, some of the data in flood estimation may be deceiving. Highly turbid waters, such as those that occur during flooding, can be easily mistaken for bare soil. The use of additional information on surrounding ground cover or vegetation in conjunction with the Landsat images may provide a better basis for flood level monitoring and estimation.

1.2 Surface Characteristics

Remote sensing provides a basis for extensive research in the identification and magnitude assessment of surface characteristics. Landsat images can be used to distinguish not only the type of vegetation covering the earth, but also the vegetation's health, what type of soil it's growing in, and of course, how much land each type of ground cover encompasses.

1.2.1 Vegetation

By comparing spectral patterns in Landsat imagery, image processors can easily detect live vegetation. Chlorophyll causes the plants to absorb strongly in the blue and red wavelengths. The biomass reflects strongly in the near infa-red region due to its cellular structure [1]. The feasibility of identifying specific crops in various fields of over 25 acres was demonstrated in a NASA study sited by Striffler. This case differentiated corn, alfalfa, and soybeans in South Dakota; wheat in Kansas; and various field and vegetable crops in California with an accuracy of 90 % or better. Forests and other vegetative land cover including grassland, brushland, deciduous forest (aspen), coniferous forest, and alpine tundra were identified with an accuracy ranging from 88 - 93 % in studies conducted in western Colorado. This study indicated that highly differing covers, such as grassland and forest, could be distinguished with confidence. Less certainty could be expected in the identification of similar features such as grassland and brushland [1].

1.2.2 Geological Applications

Remote sensing can aid in the identification of non-vegetative ground cover including minerals and soil types. Landsat images provide a large scale picture of the region of study.

By comparing spectral emittance in certain bands, numerous studies have used Landsat imagery to perform geological discrimination between even similar appearing rock formations. Accurate geological maps have been prepared solely from the imagery data in areas of sparse vegetation cover [2].

In addition to differentiating the geology of an area, identifying specific soil and soil moisture content can also be monitored using remote sensing. Studies demonstrated the effectiveness of comparing the relationship between soil spectral reflectance and soil moisture content for assessing the moisture content of soils [1]. Other effective techniques capitalize on the change in thermal properties of soils in the presence of moisture. However, the moisture content of soils with vegetative canopies are not completely reliable due to the reflectance and emittance of the canopy [1].

1.2.3 Impervious Land Cover

Due to the low level of resolution obtained through Landsat images, an accurate assessment of individual homes, buildings, and roads is impossible. By delineating urban boundaries and performing different estimations on areas contained within the urban boundary and without, studies accurately estimated the impervious land cover of 140 watersheds in the Washington D. C. area. However, these techniques were found to be less reliable in conjunction with small towns [1]. The mesh of lawns, homes, driveways, and streets results in a region which may be discernable as single family homes. The further the houses are from each other, the more easily these areas can be confused with surrounding ground cover. In areas where a class is accurately defined as single family homes, urban, etc… an estimate of

impervious land cover can be generated using a generic ratio of the impervious land cover to open land cover in such areas.

2.0 Remote Sensing

Remote sensing encompasses any type of study that gathers information about a source without coming into physical contact with it. Remote sensing studies range from simple sight or hearing observations, to astronomy, to the use of satellites for gathering information about weather patterns or the surface of the Earth. This paper concentrates on remote sensing studies involving satellite imagery of the Earth's surface.

2.1 A Brief History of Remote Sensing

Arial remote sensing began with the first known aerial photograph, taken by Parisian photographer, Gaspard Felix Tourachon (Nadar). Nadar captured Bievre, France on film from a balloon at the height of 80 meters. Kites began obtaining meteorological data via photographs around 1882. In 1909, the first aerial motion pictures were taken from one of the Wright brothers' planes. From there, remote sensing moved beyond the earth's atmosphere.

Remote sensing from space evolved from cameras launched on rockets at the beginning of the century, to pictures taken from manned mission, to satellite systems dedicated to capturing images of the Earth and its atmosphere. One of the largest providers of imagery data today is the Landsat satellite system. This program began as the Earth Resources Technology Satellite, launched on July 23, 1972, and provides repetitive monitoring of earth resources. Through space programs such as Landsat, the remote sensing

field has grown into a science capable of serving military, meteorological, agricultural, geological, environmental, and civic planning programs.

3.0 Energy Propagation Principles

Electromagnetic waves contain a sinusoidal electric wave and a magnetic wave which oscillate in phase, at right angles, both perpendicular to the direction of wave propagation. Each wave travels at the speed of light (c) and is characterized by a specific wavelength (λ) and frequency (ν). The distance from one peak to the next is the wavelength, and the number of peaks which pass a fixed point per unit time is the frequency. Frequency, wavelength, and speed are related by the equation

$$c = \nu\lambda \quad \text{(m/s)} \qquad (3.1)$$

where c is approximated as 3×10^8 m/sec. Wavelengths are generally categorized for remote sensing applications in terms of μm, or 1×10^{-6} m.

Considering only passive detection systems, the portion of the electromagnetic (EM) spectrum available for image capture ranges from 0.4 to 15 μm. This constitutes a very narrow window of the entire EM spectrum. This window includes the visible spectrum (0.4-0.7 μm), as well as infa-red (0.7-15 μm). Figure 3.1 illustrates the electromagnetic spectrum, highlighting the portion usable in remote sensing.

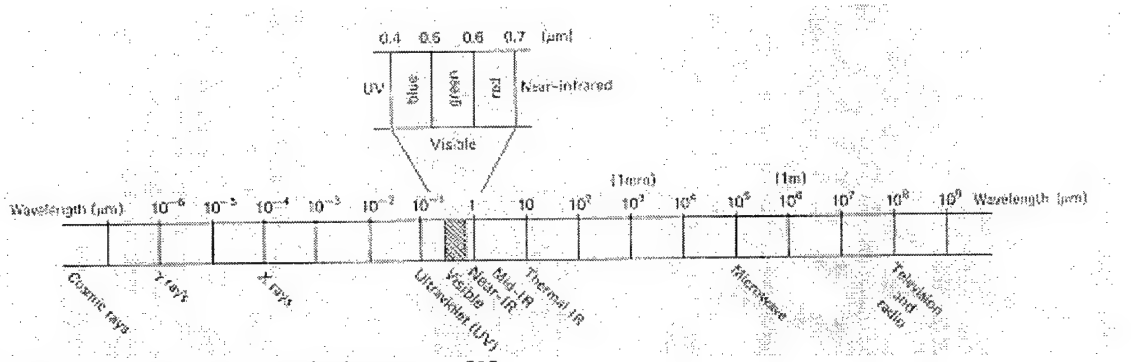


Figure 3.1: Electromagnetic Spectrum [3]

All matter emits electromagnetic radiation. The total exitance over all wavelengths can be characterized by the *Stefan-Boltzmann law,*

$$M=sT^4 \qquad (3.2)$$

Where

M = total radiant exitance from the surface of a material (Wm^{-2})
s = *Stefan-Boltzmann constant,* 5.6697×10^{-8} ($Wm^{-2}K^{-4}$)
T = absolute temperature of the emitting material (K)

However, this equation only applies to blackbodies. A blackbody behaves as a perfect radiator. It totally absorbs and emits all energy incident upon it. The wavelength at which peak blackbody exitance occurs, the dominant wavelength, can be calculated from its temperature through *Wein's displacement law,*

$$\lambda_{max} = A/T \qquad (3.3)$$

Where

λ_{max} = wavelength of maximum spectral radiant exitance (μm)
A = *Wein displacement constant,* 2898 (μm K)
T = temperature (K)

The sun acts as a blackbody radiating at about 6000 K. This corresponds to a dominant wavelength of approximately 0.5 μm. This correlates with the center of the visible spectrum. However, objects with temperatures near the earth's ambient temperature, about

300 K, produce a dominant wavelength of approximately 9.7 μm. This corresponds with the general distinction of infrared, or thermal, energy [4]. These relationships are illustrated more clearly in figure 3.2.

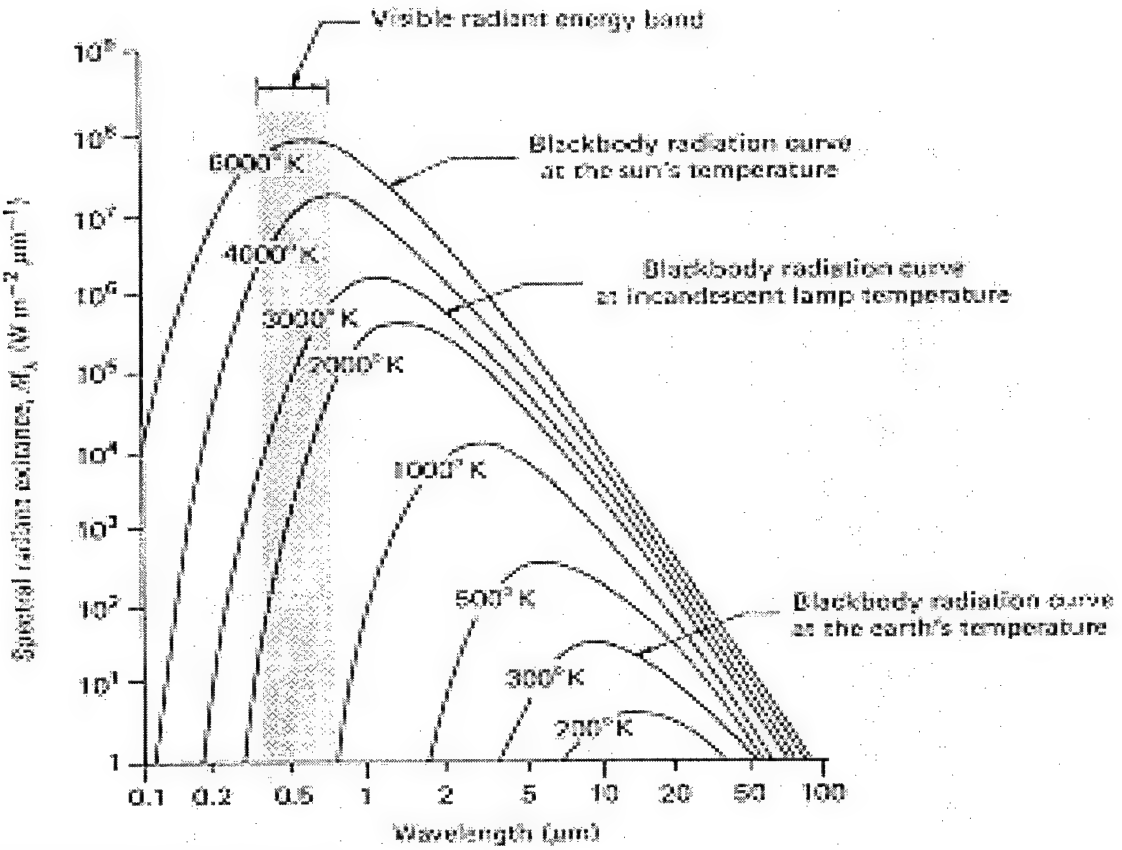


Figure 3.2: Blackbody Curves [3]

Why do we see images in a spectral zone that they are not emitting strongly from? This is because the light we see is reflected off the object and originates from the sun. Therefore, as a general rule, below wavelengths of approximately 3 μm, reflected energy dominates. Above 3 μm, the energy is assumed to be emitted [4].

The atmosphere's composition also limits the amount of reflected or radiated energy that can reach a sensor. Water vapor, carbon dioxide, and ozone act as the primary absorbers

[3]. This absorption creates windows in the EM spectrum for remote sensing energy measurements. Three types of scattering also limit EM transmission through the atmosphere. Rayleigh scattering occurs when particles much smaller than the wavelength bend it, affecting primarily the shorter wavelengths, resulting in blue skies. When particles nearly equal to the wavelength in diameter, such as water vapor and dust, interact, Mie scattering occurs. Non-selective scattering affects all remote sensing wavelengths because particles are much larger than the wavelengths, such as water droplets. Fog and clouds appear white as a result.

Figure 3.3 combines the effects of atmospheric transmission, solar irradiance, and ambient earth temperature exitance. Note that the visible spectrum occurs in wavelengths of high atmospheric transmission. However, not all of the infrared portion of the spectrum is included. Therefore, bands of measurement through the infrared spectrum coincide with the windows available. The thermal band corresponds with the window around 10 μm which also corresponds to the dominant wavelength of earth temperature and exitance.

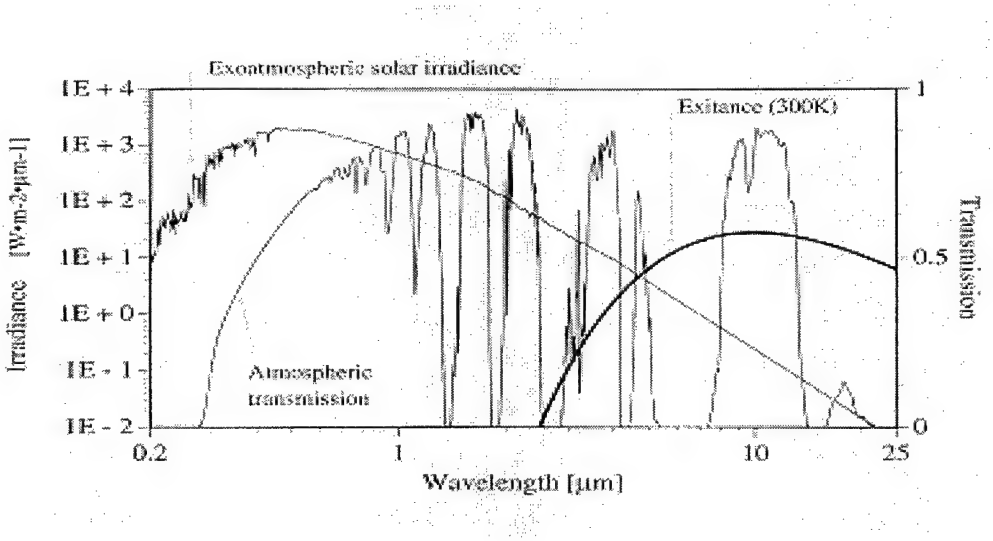


Figure 3.3: Exoatmospheric Solar Irradiance, Atmospheric Transmission, and Exitance versus Wavelength [4]

3.1 Spectral Response Curves

Every object reflects and emits energy in a specific way. This is how the human eye can distinguish between objects and colors. In each spectral band, a certain amount of energy is reflected or emitted from each object. When graphed over the entire spectrum, these energy responses result in a spectral response curve, as illustrated in figure 3.4.

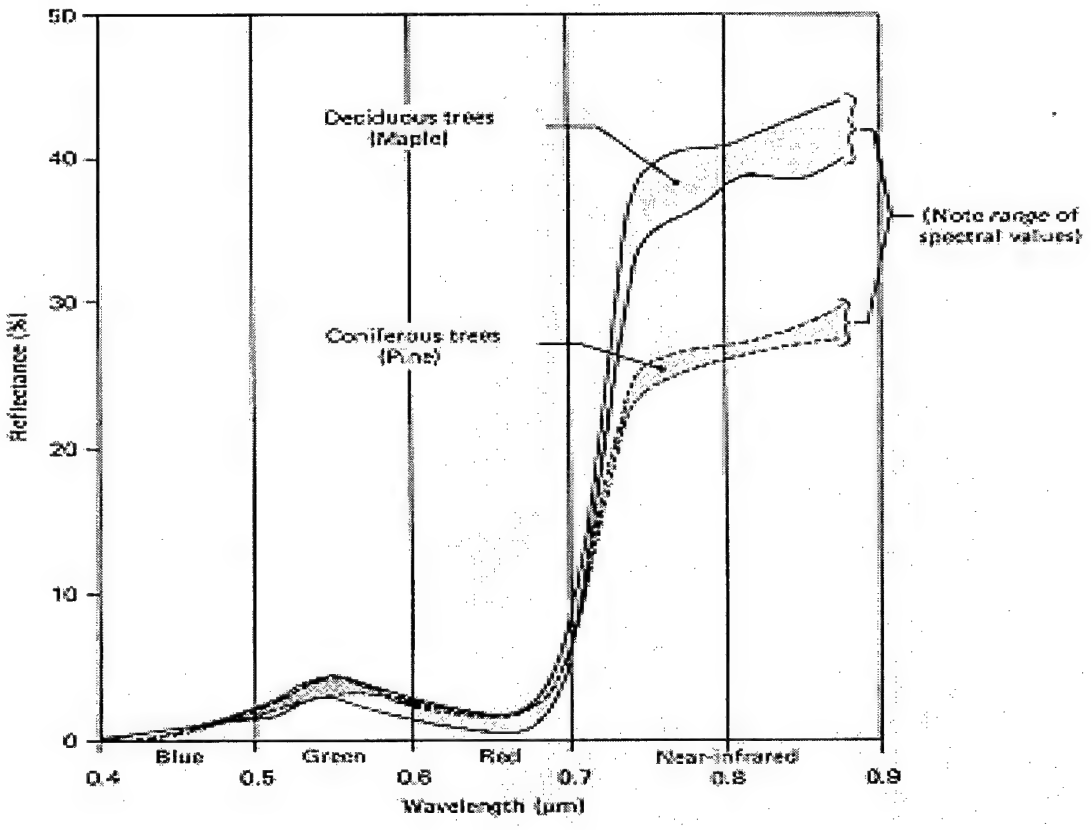


Figure 3.4: Spectral Response curves for coniferous and deciduous trees [3]

Figure 3.4 demonstrates the use of spectral response curves for classification. In some of the bands, particularly those between 0.4 and 0.7 μm, both coniferous and deciduous trees give the same response patterns. When viewing those bands, no difference would be noted

between the types of trees. However, when viewing bands containing wavelengths between 0.7 and 0.9 μm, a drastic difference is noted between the responses of the two types of trees.

The curves can help classification in two ways. First, the curves can be viewed before attempting to differentiate between objects. Then, the classifier knows what bands to look at to distinguish between the objects. For this example, it would be the bands containing 0.7-0.9 μm. Otherwise, if a classifier finds an odd response when scrolling through the available bands, he or she can match spectral response curves to known response curves for a distinct classification of objects. In this case, the classifier may not be aware that more than one type of tree is contained in the sample area, but when differing responses are noted, the classifier finds that indeed, the response curves correlate to both deciduous and coniferous trees.

For vegetation studies, spectral response curves serve only as a guide to what bands are most useful in classification, and the relative response patterns exhibited by various species. Due to seasonal changes and other variables in the growth cycle such as water availability, pests, etc., definitive spectral response patterns are not generally consulted for vegetation ground cover. However, relative response patterns generated for the particular study site usually provide ample information for separation of vegetation classes.

4.0 Image Capture

Remote sensing offers two types of image capture, analog and digital. Analog provides a continuous spectrum of values for light intensities contained in an image. Digital capture employs a stepwise recording of radiation through the use of charged couple devices (CCD's). Analog can therefore provide a more accurate display of the image; however, it cannot be manipulated to enhance features. Digital capture may lose some of the subtleties in

intensity variations due to its stepwise nature, but it is capable of providing a broader range of brightness sensitivity than analog capture, as illustrated in figure 4.1. In addition, digital capture enables the user to manipulate and interpret images with flexibility that far exceeds analog capture's abilities. Most satellite systems, including Landsat, employ digital capture of images.

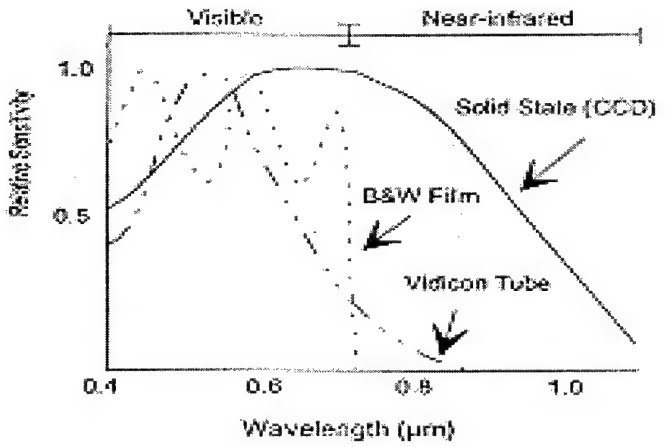


Figure 4.1: Sensitivity comparison of image capture options [5]

4.1 The Landsat System

Landsat-4 and -5 satellites weight approximately 2000 kg. They include solar panels mounted on one side, a multi-spectral scanner, thematic mapper, X-band and S-band antennas, and a high gain antenna mounted on a boom. The X-band and S-band antennas provide direct data transfer when necessary. The high gain antenna enables the satellite to relay data to ground sites through the geosynchronous communication satellite network, Tracking and Data Relay Satellite System (TDRSS) [3]. Figure 4.2 illustrates the Landsat-4 and –5 configurations.

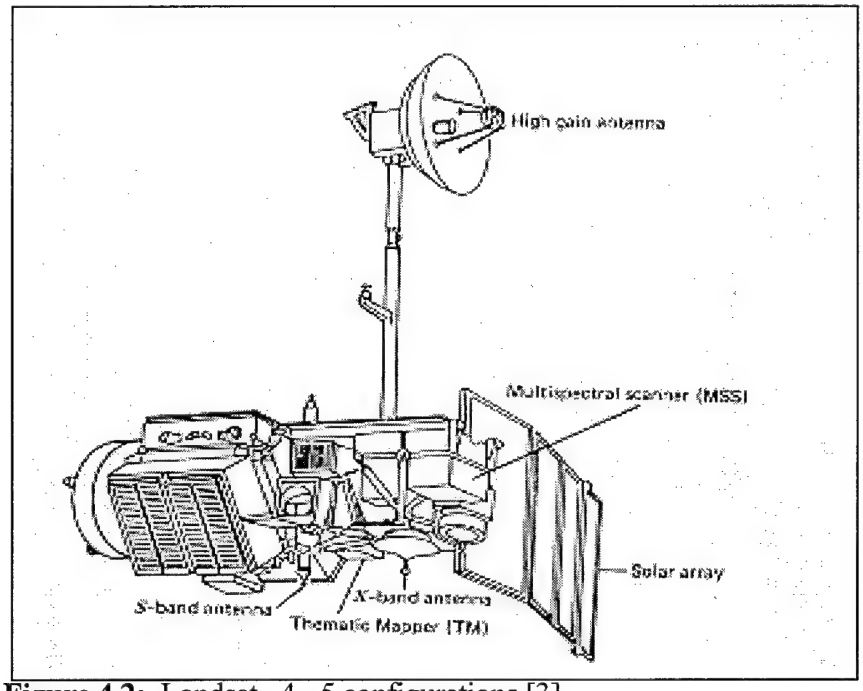


Figure 4.2: Landsat –4, -5 configurations [3]

4.1.1 Orbit and Constellation

Landsat's current satellites follow repetitive, circular, sun-synchronous, near-polar orbits at a height of 705 km above the earth. This lower orbit makes the satellites potentially achievable by the space shuttle, as well as improving their resolution over the first three Landsat satellites. Landsat-1-3 orbited at heights of 900 km [3].

Landsat orbits at an inclination of 98.2 degrees. It crosses the equator at 9:45 A.M. local sun time every day (sun synchronous). Each satellite takes approximately 99 minutes to complete an orbit. This corresponds with a 16 day ground track repeat. Landsat 4 and 5 are eight days out of phase, so that an 8 day repeat cycle is established when both are operational. The distance between consecutive day's ground tracks is approximately 2752 km at the equator. Adjacent ground tracks are taken every 7^{th} day [3]. The ground track scenario is depicted in figure 4.3.

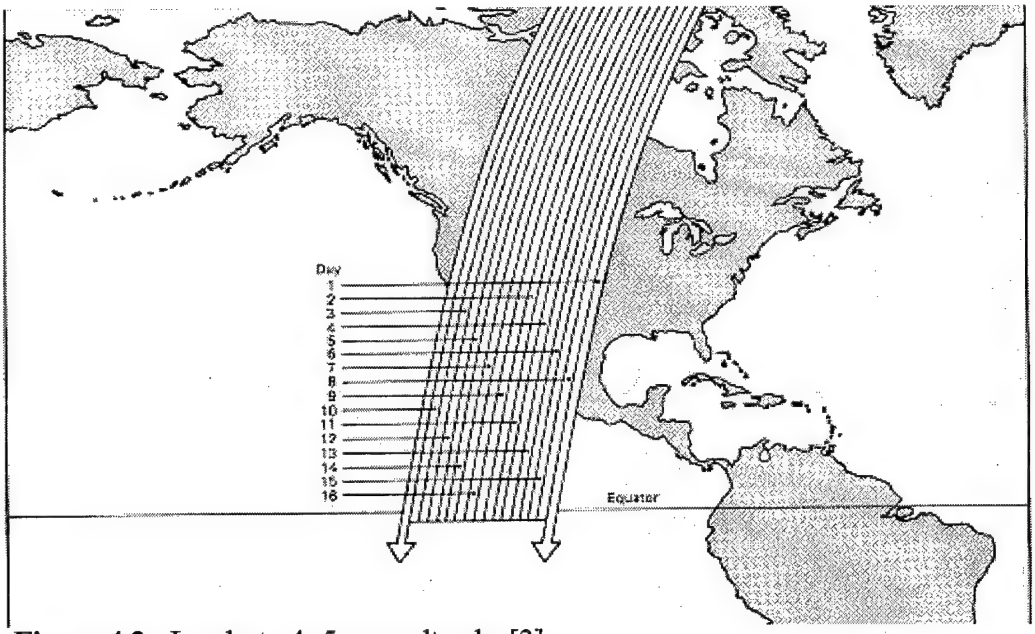


Figure 4.3: Landsat –4,-5 groundtracks [3]

4.1.2 Sensors

Landsat-4 and –5 include both the multi spectral scanner (MSS) and thematic mapper (TM). Both of these systems employ across-track scanning [5]. Energy from the field of view is split into the spectral components desired for collection. The visible light is separated via a prism, whereas a dichroic grating separates the thermal component. The resulting narrow bands of energy are then projected onto an array of detectors [6]. The voltage at each detector is sampled periodically (every 9.95 μs for the MSS), and converted to a gray scale value. This value corresponds to the darkness of one pixel in the final image and is stored using either 6 or 7 bits (64/128 values) for the MSS, or 8 bits (256 values) for the TM. Table 4 delineates the spectral bands detected by each system.

Sensor	Band	Waveband (μm)	Pixel (m)	Levels
MSS	1	0.5-0.6	79	128
	2	0.6-0.7	79	128
	3	0.7-0.8	79	128
	4	0.8-1.1	79	64/128
TM	1	0.45-0.52	30	256
	2	0.52-6.0	30	256
	3	0.63-0.69	30	256
	4	0.76-0.90	30	256
	5	1.55-1.75	30	256
	6	10.4-12.5	30	256
	7	2.08-2.35	30	256

Table 4: Characteristics of the MSS and TM [5]

4.1.2.1 Multispectral Scanner

The Landsat program has contained a multispectral scanner in every spacecraft. It provides data over a range of spectral windows to help generate spectral signatures for classification and comparison of features contained in the images. Landsat MSS data constitutes the most comprehensive remote sensing database in the world. The MSS captures information in four bands including, 1. 0.5-0.6 μm, 2. 0.6-0.7 μm, 3. 0.7-0.8 μm, and 4. 0.8-1.1 μm. The MSS scans the 185 swath from west-to-east by oscillating a small mirror over a 14.92 degree total field of view. The mirror oscillates once every 33 msec [3]. The MSS acquires 6 lines in every scan. Therefore, ground coverage occurs at 1/6 of the single line scan rate. Since six lines are scanned, six sensors must be present for image capture in each band

because every line requires a sensor. The MSS captures 4 bands with 6 sensors in each band, so it contains 24 sensors. Figure 4.4 illustrates the Landsat multispectral scanner system.

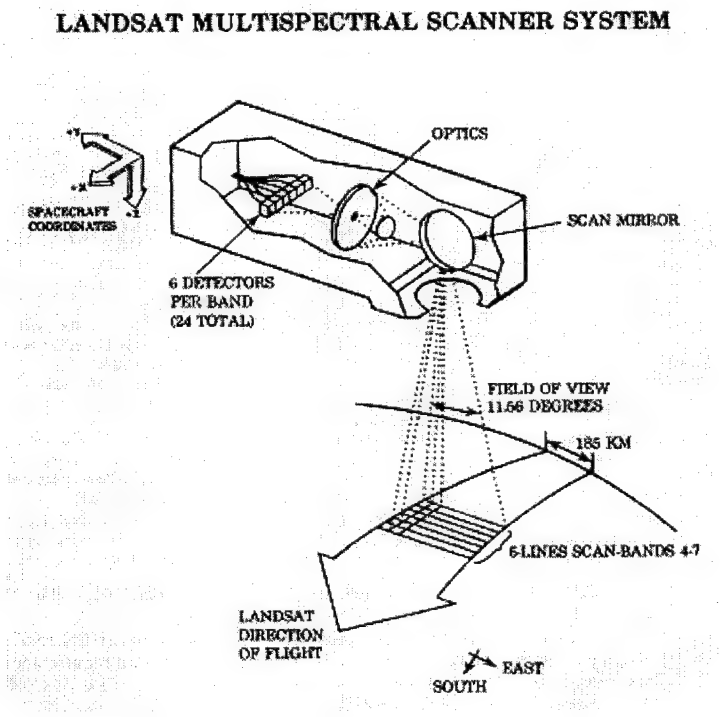


Figure 4.4: Multispectral Scanner System [7]

The MSS samples each sensor's voltages as it sweeps across the swath and converts the voltage into a value between 0 and 127 for digital processing and display. This sampling process generates approximately 3240 pixels per line. Dividing the total swath length of 185 km by the sampling rate in the time for each line to be scanned results in an actual frame length of 56 m. However, the brightness value for each pixel is derived from an 82 meter square, equal to the swath width [3]. The resulting matrix is composed of 56 x 82 m^2 rectangles. The MSS instantaneous field of view is equal to the sample area, 82 x 82 m^2, or 6724 m^2 [7].

4.1.2.2 Thematic Mapper

The thematic mapper was added to Landsat –4 and –5 as an improvement on the MSS image acquisition. The TM obtains data for 7 spectral bands, rather than four. The extra bands improve spectral differentiability of major earth features and cover a broader range of the EM spectrum than the MSS. These bands include, 1. 0.45-0.52 μm, 2. 0.52-0.6 μm, 3. 0.63-0.69 μm, 4. 0.76-0.90 μm, 5. 1.55-1.75 μm, 6. 10.4-12.5 μm, and 7. 2.08-2.35 μm. They cover the visible spectrum, plus another blue band, near infrared, mid-infrared, and thermal portions of the spectrum [5].

The TM covers the same swath width as the MSS, 185 km. However, it acquires data when scanning in both the west-to-east direction as well as the east-to-west direction. This reduces the scan rate and increases the dwell time for each detector. It only completes 7 cycles in one second over the total field of view, 15.4 degrees. This reduces the acceleration of the mirror to improve geometric integrity and signal to noise performance for the system [3]. TM data are collected using a 30 m ground resolution cell for non-thermal, and 120 m for thermal bands. This reduces the ground cell area by 7 from that obtained with the MSS.

The TM employs 16 detectors for each non-thermal band and 4 for the thermal band. This corresponds to 100 detectors compared with the MSS's 24. Both sets of detectors are calibrated using three tungsten filament lamps, a blackbody, and a pivot mounted shutter. The shutter directs the lamps' light to the non-thermal band (1-4) detectors, while a mirror on the shutter directs the blackbody energy to the thermal band (5-7) detectors. Detectors for bands 1-4 are located in the primary focal plane, and detectors for bands 5-7 are located on a cooled second focal plane. Every detector views a different area on the ground due to the spatial

separation of the detectors on the two focal planes [3]. Figure 4.5 illustrates the TM optical path and assembly.

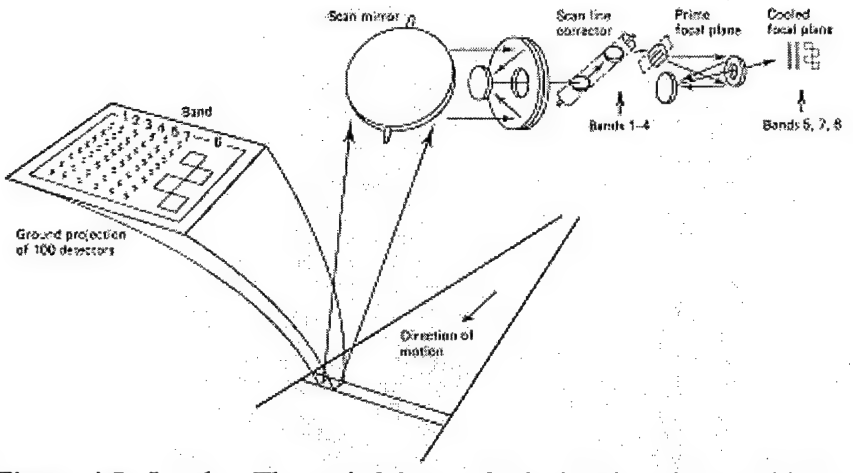


Figure 4.5: Landsat Thematic Mapper Optical Path and Assembly [3]

Spacecraft motion and spatial separation necessitate corrections based on time for band-to-band correlation. A scan angle monitor generates signals indicating the mirror's angular position as a function of time, called scan mirror correction data. This data is transmitted to the ground for geometric image correction and to the scan line corrector. The scan line corrector rotates the TM line-of-sight backwards along the satellite ground track to compensate for the forward motion of the spacecraft. This produces straight scan lines that do not overlap or underlap [3].

5.0 Image Interpretation and Analysis

Because this paper focuses on Landsat imagery, only digital image processing will be discussed in this portion. However, an important step in interpretation and analysis of the image involves gathering sufficient ground truth. This data can come from a variety of sources and is correlated to the image to enable accurate classification. Other steps in the

classification process include image enhancement, an applied classification method, smoothing, and error assessment. This section explains each step and its applications using Idrisi software and Thematic Mapper data gathered for the Howe Hill area in Massachusetts.

5.1 Rectification and Enhancement

Due to the southerly motion of the satellite, and the easterly rotation of the earth, the actual image captured by Landsat satellites looks more like a parallelogram than a rectangle. Each successive line captured falls slightly to the west of the previous line. This result is a systematic and predicable error. It can be easily rectified through comparison of the image to USGS maps.

Other systematic errors are imposed on the image by the detectors. Each detector has a specific range of sensitivity calibrated using onboard calibration lamps [3]. The detector senses a certain radiance as its 0 digital number (DN) value, and a different radiance as its 255 DN value. The range in between corresponds to a linear relationship. These min and max values are posted in the header of each image, so absolute radiance values for any given image can be calculated from the equation:

$$L = \left(\frac{LMAX - LMIN}{255} \right) \cdot DN + LMIN \qquad (5.1)$$

The relative radiances generated by Landsat are sufficient for the purposes of this paper, so absolute radiance values were not generated.

Landsat images display energy as a digital number (DN) for every pixel in the scene in each band measured. Often, the range of values does not encompass the entire spectrum of available values (0-255). To enhance different features, a contrast stretch

may be applied. In this case, the values are spread to use the entire spectrum. The lowest value becomes 0, and the highest equals 255 regardless of the original DN. This stretch enhances variations in bright or dark areas which may not be apparent in the original spectrum. This type of stretching can be applied even if the DN values span the entire range. In this case, a certain percentage of DN values can be truncated from each end of the scale, and the linear stretch is then applied to the remaining DN range. This enhances differences in energetic response for easier identification of features.

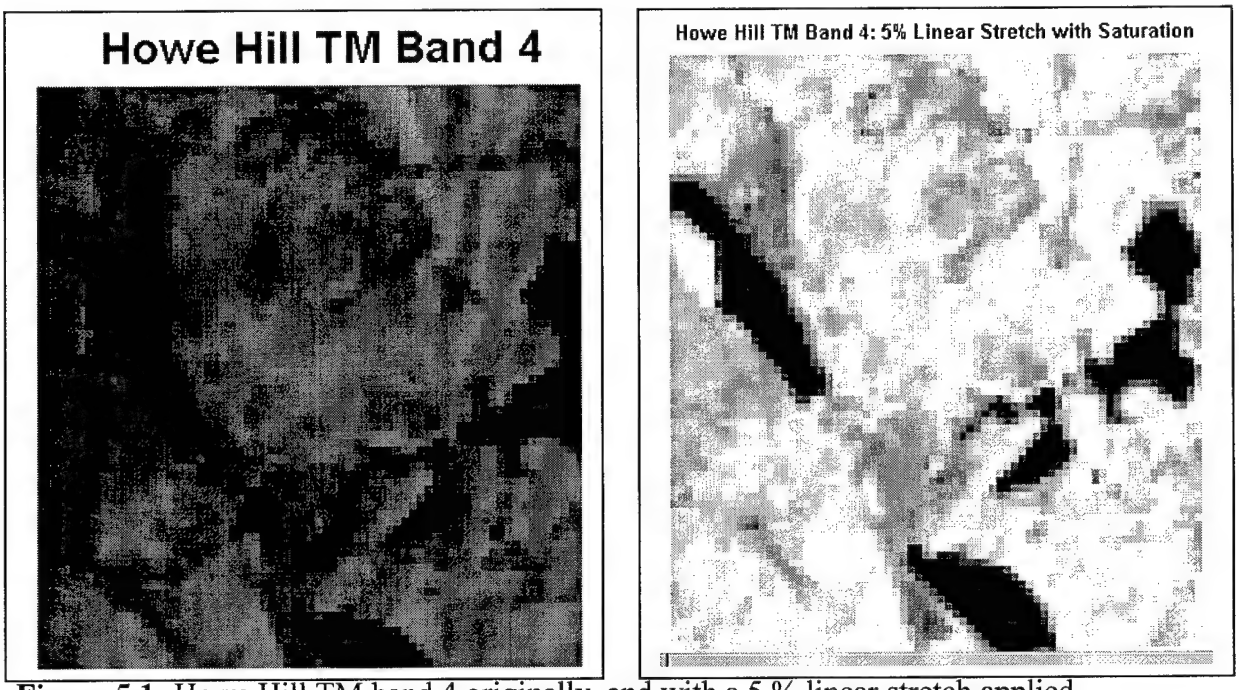


Figure 5.1: Howe Hill TM band 4 originally, and with a 5 % linear stretch applied.

Figure 5.1 demonstrates the application of contrast stretching using a 5% truncation. Appendix A contains histograms of the DN values spread and occurrences for each image. Note that the DN values range from 0 to 190 in the original histogram. 5% of the DN values are truncated from each side of the histogram. The remaining DN values are then stretched to cover the entire spectrum, 0 to 255. The relative differentiation of this stretch provides clear distinctions between water and land. Water appears very dark in both images, but is in greater

contrast in the stretched image. Linear stretching obviously aides in visual analysis, but is also a useful tool in computer classification.

5.2 Classification

Digital classification methods fall into either of two categories, supervised and unsupervised classification. Both methods employ the use of multi-spectral data. These data are combined to form spectral response patterns that are then used to separate and classify each pixel into a specific group. Supervised classification specifies the number of groups the data is to be divided into, while unsupervised classification allows the data to fall into groups defined only by the spectral response patterns, not a priori knowledge of site content.

5.3 Supervised Classification

Supervised classification using Idrisi, can accommodate the use of up to seven spectral bands, the number of bands Landsat's Thematic Mapper captures. The basic sequence of operations for supervised classification follows:

1. Define Training Sites
2. Extract Signatures
3. Classify the Image
4. Post-classification Smoothing
5. Accuracy Assessment

5.3.1 Define Training Sites

A single band, or composite of up to three bands can be used to define training sites in Idrisi. For this demonstration, the infrared image, TM band 4, was used. A land use map of Howe Hill (figure 5.2) delineated appropriate training sites for supervised classification.

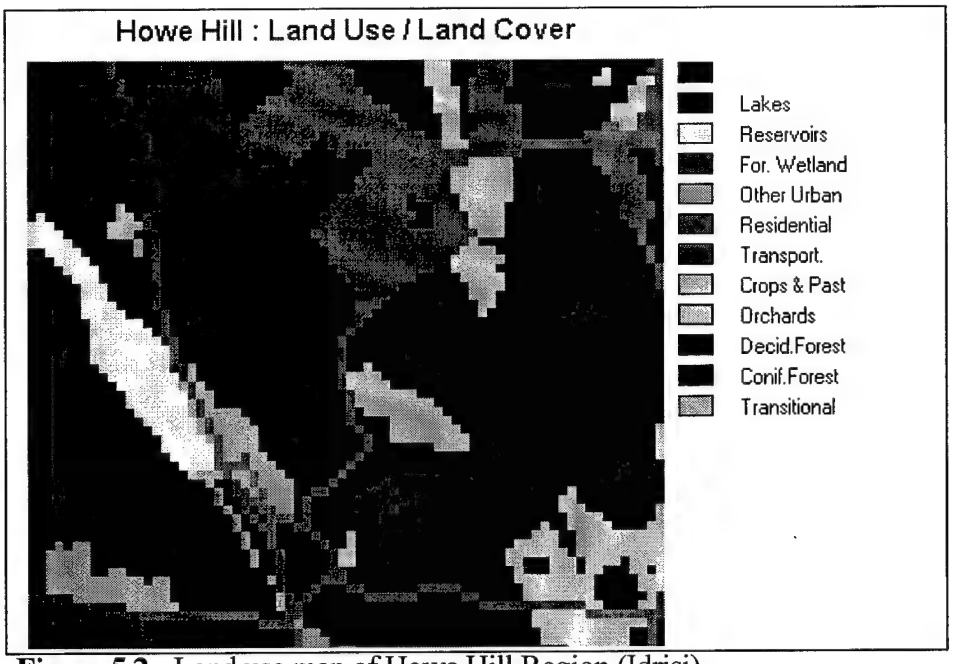


Figure 5.2: Land use map of Howe Hill Region (Idrisi)

Figure 5.2 is an example of ground truth. Ground truth may be gathered through samples for a specific study, through examining geological, agricultural, and hydrological studies of the site, or by employing Geographical Information Systems (GIS). GIS's contain layered information files on sites including information such as mineral content, ground cover, etc. This site was provided by Idrisi as a training exercise. In order to locate appropriate training sites in the field, areas of at least several pixels must be identified and located. These sites must be homogeneous and can be located using GPS receivers and/or a map. Obviously, the larger the area, the higher the potential for larger homogeneous areas for training sites which subsequently improve classification. In general, at least 10 times as many pixels must be characterized as there are bands in the image to classify. Therefore, at least 70 pixels must be sampled for a Landsat TM image [8].

5.3.2 Extract Signatures

Idrisi now extracts signatures from the training sites in the image. The signatures correspond to the response in each band for the area contained in the training site. Idrisi

generates a characterization of the each class defined in the training set. In this step, I chose to use only 5 signatures even though 11 classes are defined by the land use map. The classes were derived from initial analysis, and the ability to define adequate training sites. The classes include water, coniferous trees, urban, agriculture, and deciduous trees.

5.3.3 Classify the Image

Now, Idrisi applies the characterization obtained through the training site to each pixel in the image. Idrisi offers two types of classification, soft and hard. Soft classification allows for the mixing of classes within each pixel. For example, a pixel may contain 20% conifers, and 80% deciduous trees. The soft classifier "expresses the degree to which a pixel belongs to each of the classes being considered" [8]. Hard classifiers yield a definite association with one of the defined classes. If 46% of a pixel belongs to coniferous and 54% deciduous, the pixel is classified as deciduous. I chose a hard classifier for this example since the ground truth was defined using whole pixels.

Idrisi offers several options within the hard classifier category. The classifier compares the response zones of each class in every band to the unknown pixel. MINDIST classifies the unknown pixel by the minimum distance to the mean of each training class. PIPED generates a parallelepided region for each class. Whatever parallelepided the unknown pixel falls into classifies the pixel. However, in cases of parallelepiped overlap, the classification is arbitrary. The classifier MAXLIKE is based on Bayesian probability theory. MAXLIKE uses the mean and variance/covariance data of the signatures to generate elliptical zones of probability for each class. The unknown pixel falls within a zone for every class. The class with the highest probability zone for that pixel defines the pixels classification. I

chose MAXLIKE as the classifier for this data because it produces the best results although it is slightly more time consuming [8].

5.3.4 Post-Classification Smoothing

Once classification is complete, the resulting response histograms for each band can be viewed. Ideally, these histograms result in a bell shaped curve centered on a DN of greatest frequency. Therefore, if outliers exist, these can be severed from the data set to provide a better curve and resulting better data sets as illustrated in figure 5.3. The left image histogram contains outliers which are severed in the right image histogram.

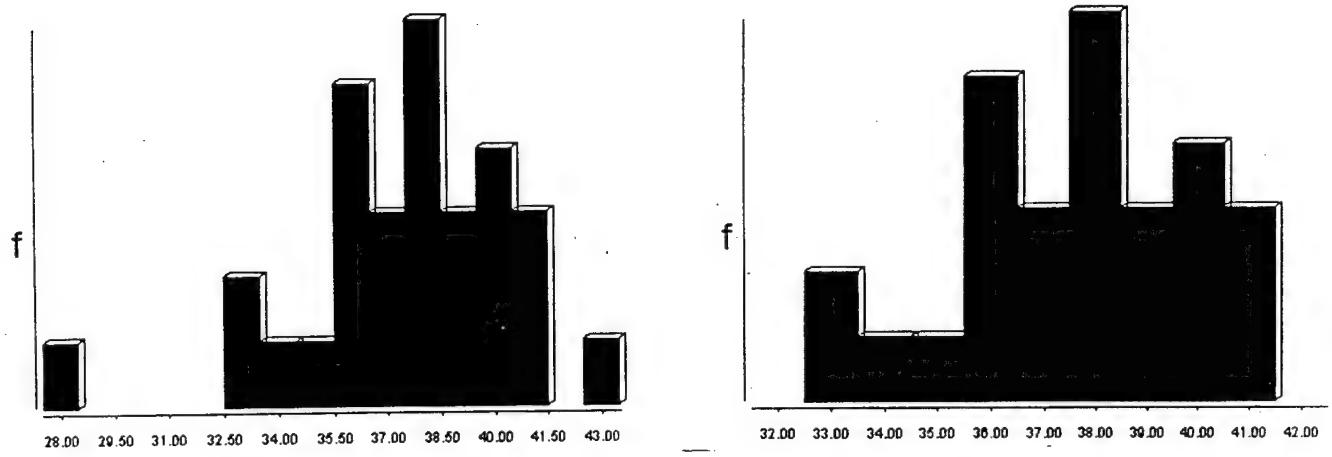


Figure 5.3: Conifers TM band 5 with outliers, and without

If the data set exhibits multiple peaks, this indicates that several types of classes are merged into one. Either reclassification can result, or the histograms can be compared to the other existing classes for a match. If a match occurs, this indicates that the training site contained both its own type of cover, and the matching cover. This type of distortion is exhibited in figure 5.4. Notice how the deciduous peak is contained in the agriculture

histogram. In this case, the deciduous peak can be separated from the agriculture class, resulting in a new agriculture class exhibited in figure 5.5.

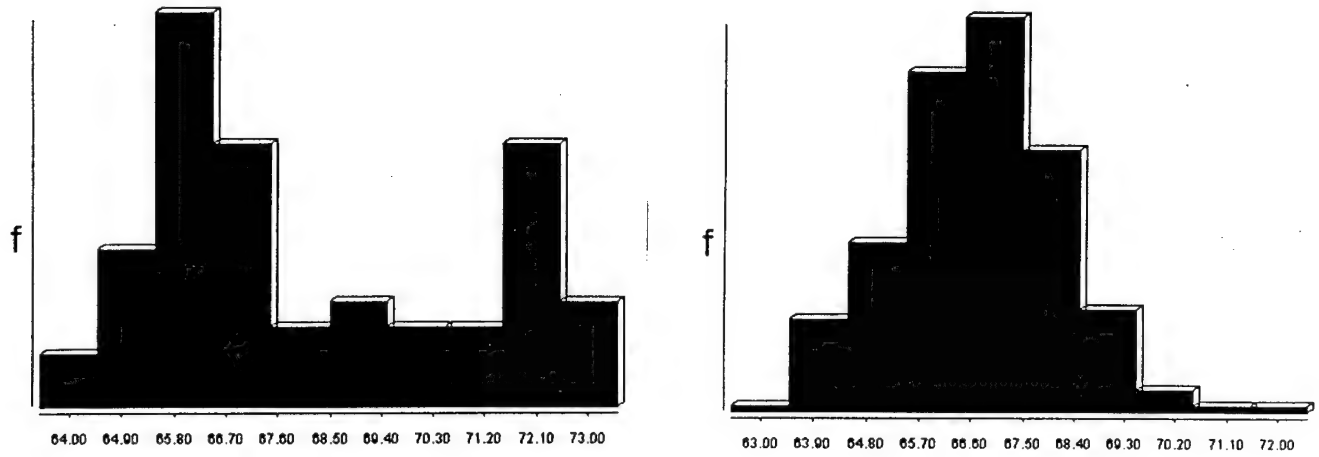


Figure 5.4: Histograms of agriculture TM band1 versus Deciduous TM band 1

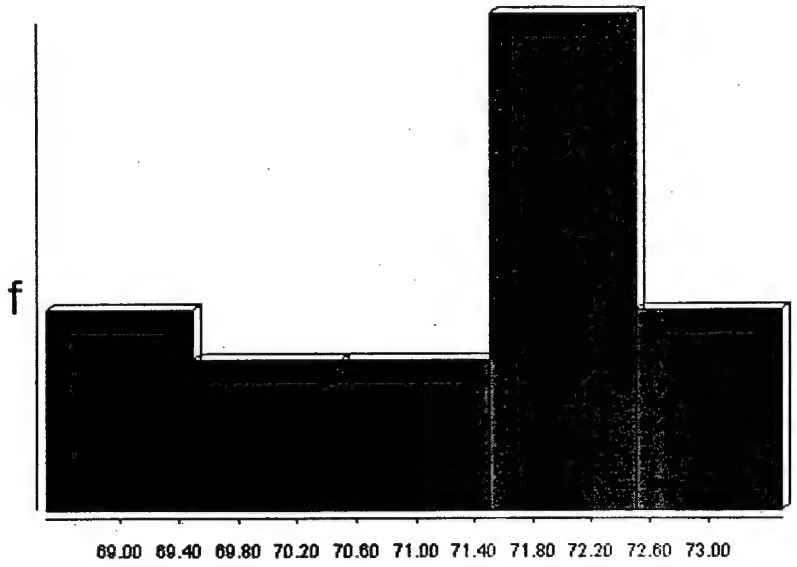


Figure 5.5: New agriculture histogram for TM band 1

5.3.5 Accuracy Assessment

I measured the accuracy of the MAXLIKE classification using ERRMAT in Idrisi. Essentially, ERRMAT produces what is called an error matrix. The error matrix compares the

relationship between known reference data, and the classification data generated by Idrisi on a category by category basis [3]. The known classification (columns) versus the classifier based pixels (rows) generates the matrix. The error matrix generates errors of omission (exclusion) as well as errors of commission (inclusion). The overall accuracy and the accuracy of each class can also be determined (see Appendix B). Normally, only select reference areas are used for this analysis. The accuracy assessment only indicates how well the classifier performs in those areas and cannot necessarily be applied as an accuracy assessment of the entire area [3]. However, since a land use map for the entire Howe Hill area exists, the error matrix generated by ERRMAT does apply to the entire image.

5.4 Unsupervised Classification

Idrisi software utilizes a composite image for unsupervised classification. The image can contain only three spectral bands. I chose LANDSAT bands 3, 4, and 5, which correspond to the visible red band, near-infrared band, and middle infrared band, respectively, and provide the most information [8]. I then applied a 1 % saturated linear stretch as suggested in the Idrisi User's Guide. The sequence of operations for unsupervised classification follows:

1. Cluster classification
2. Post-classification smoothing
3. Accuracy Assessment

Post-classification smoothing and accuracy assessment follow the same format as supervised classification, so this section only further examines cluster classification.

The CLUSTER option was applied to indicate every class Idrisi could separate, which generated 14 classes. However, these classes were not well defined. Therefore, I examined the histogram of the composite DN values (figure 5.6).

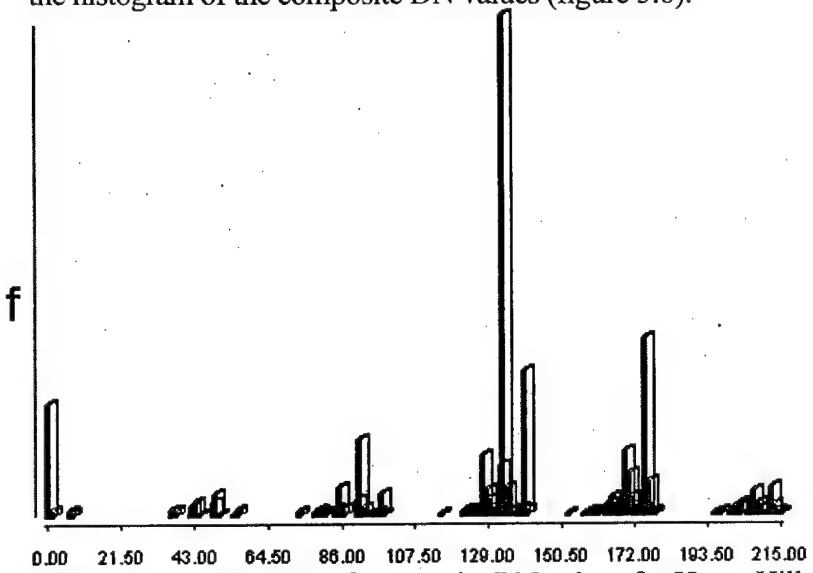


Figure 5.6: Histogram of composite DN values for Howe Hill, TM bands 3, 4, 5

Figure 5.6 clearly indicates the presence of 6 classes. I then ran CLUSTER with a maximum of six classes. These classes still indicated some ambiguity, so two of the classes were subsequently combined as urban.

5.5 Combination Classification

Combination classification merges unsupervised classification with supervised classification. First, I ran CLUSTER on the composite image to delineate the 5 classes, then I applied the training sites as in supervised classification, following the class definitions on the unsupervised image. From there, smoothing and accuracy assessment followed. The steps for combination classification follow:

1. Cluster classification
2. Define training sites
3. Extract signatures

4. Classification
5. Post-classification smoothing
6. Accuracy Assessment

6.0 Results and Discussion

I generated classification scenes using supervised classification and two levels of smoothing, unsupervised classification, and combination classification using two levels of smoothing. The first level of smoothing consisted of simply cropping the outliers from the response DN histograms for each band in every class. The second level of smoothing involved cropping the histogram to form a more bell shaped curve, and separating merged classes, as described for the agriculture band above. I generated the results using the ERRMAT function in Idrisi. The error matrices can be found in Appendix B.

Classification Technique	Smoothing Level	Overall Accuracy
Supervised	None	0.6457
Supervised	1	0.6410
Supervised	2	0.6428
Unsupervised	None	0.6315
Combination	None	0.5782
Combination	1	0.5719
Combination	2	0.5696

Table 6.1: Overall accuracy of applied classification techniques

Supervised classification surpassed the performance of combination classification by seven percent. Unsupervised classification performed comparably to supervised classification.

This indicates that the spectral responses of each class were fairly well defined. However, an accuracy of 64 % does not indicate a very good classifier. Generally, classifications of 80 % accuracy or better are preferred [3]. Surprisingly, overall accuracy was diminished by using post-classification smoothing techniques in both supervised and combination classification.

Combination classification provided the worst estimate. The ground truth sites I chose for this portion of the study were areas that were completely enclosed in the land use map, but appeared salt and peppered on the unsupervised classification image, with the exception of water. The choice of ground truth sites may have led to poorer accuracy in combination classification.

Simply comparing overall accuracy does not tell the whole story. Some classes achieved a high level of classification accuracy, where others performed dismally. This pattern occurred for every classification technique used. Table 6.2 displays the errors of commission and omission for unsupervised classification without smoothing. This provides an adequate model for all of the classification techniques.

Class	Description	Commission	Omission
1	Water	0.1072	0.1439
2	Conifers	0.3556	0.5276
3	Urban	0.4602	0.3593
4	Agriculture	0.7815	0.3308
5	Deciduous	0.0920	0.3711

Table 6.2: Errors of commission and omission for supervised classification without smoothing.

Errors of commission correspond to the probability that a pixel defined as a certain class will not be that class in the field. For both the water and deciduous classes, pixels designated as either class have approximately a 90 % probability of actually being that land cover at the site. Agriculture, on the other hand, has only a 22 % probability of its classification agreeing with ground truth. This means that there are many extra pixels included in the agriculture class that do not actually belong there.

Errors of omission correspond to the probability that pixels of a certain class in the field are not designated as that class during classification. Here, water still performed the best. Only 14 % of ground truth that was actually water was not included in the class. Conifers had the highest percent probability of not being classified correctly, with 53 %. The other three classes had a 33-37 % probability that their classifications would be falsely contained in another class.

These errors indicate the performance ability of the training sites. The training site for water was excellent. None of the other training sites were very homogeneous, especially agriculture. These errors stem from the fact that broad areas were defined on the land use map as deciduous, conifers, urban, and agriculture, whereas most of those regions probably contained a mix of several categories. Actually visiting the field site could indicate more appropriate training sites. Another factor is the use of hard classifiers. In cases where a variety of classes are present in a pixel, only soft classifiers will indicate this. However, the "ground truth" land use map indicated hard classification of pixels.

7.0 Conclusions

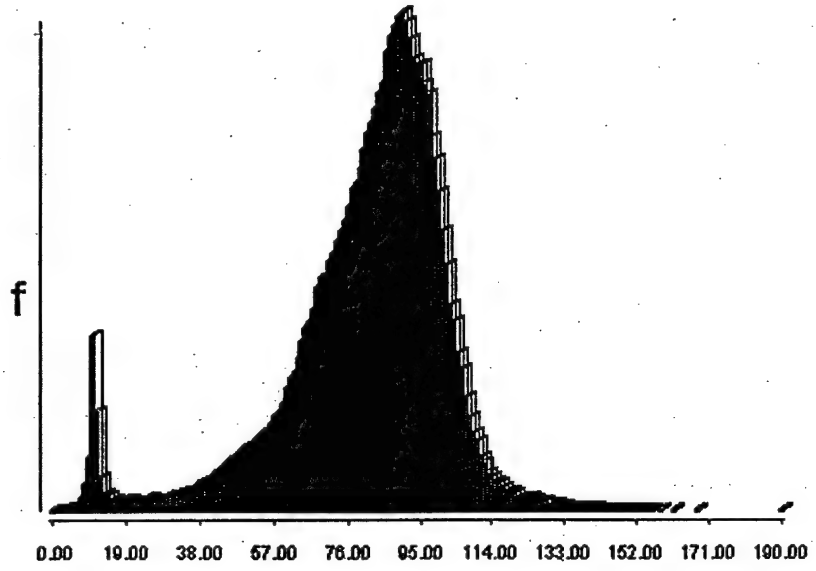
A classifier is only as accurate as the training sites. Large areas (at least 4-5 pixels) must be defined and accurately placed on the image for supervised classification or

combination classification to be effective. Unsupervised classification can give a good indication of where these training sites may be located quickly, without visiting the field site before classification. Without the proper use of training sites and ground truth, both supervised and combination classification perform with accuracies equivalent to unsupervised classification accuracy. This illustrates the fact that a strictly digital classification without proper ground truth correlation cannot provide the accuracy necessary for land use classification.

References:

1. Striffler, William D., and Diane C. Fitz. Application of Remote Sensing in Hydrology. Fort Collins, Colorado: Colorado State University, 1980.
2. Linz, Joseph Jr., and David S.Simonett Ed. Remote Sensing of Environment. Reading, Massachusetts: Addison-Wesley Publishing Company, 1976.
3. Lillesand, Thomas M., and Ralph W. Kiefer. Remote Sensing and Image Interpretation 3rd Edition. New York: John Wiley and Sons, Incorporated, 1994.
4. Schott, John R. Remote Sensing: The Image Chain Approach. New York: Oxford University Press, 1997.
5. Wilkie, David S., and John T. Finn. Remote Sensing Imagery for Natural Resources Monitoring. New York: Columbia University Press, 1996.
6. Augesteijn, Dr. Marijke F. "Remote Sensing in Space course notes." Fall, 1997.
7. Jensen, John R. Introductory Digital Image Processing, A Remote Sensing Perspective. Englewood Cliffs, New Jersey: Prentice-Hall, 1986.
8. Eastman, Ronald J. Idrisi for Windows User's Guide Version 2.0. Worchester, MA: Clark University, 1997.

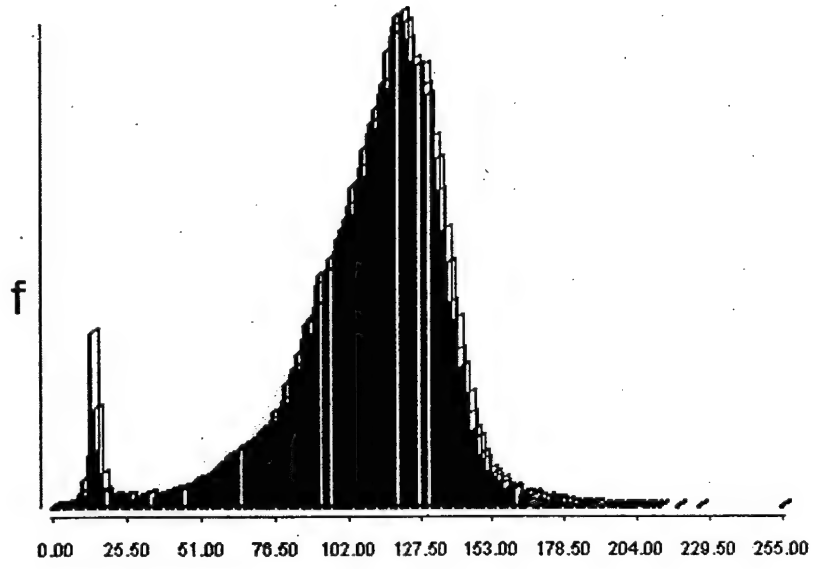
Appendix A



Histogram of
HOW87TM4

Class width :	1.0000
Display minimum :	0.0000
Display maximum :	190.0000
Actual minimum :	0.0000
Actual maximum :	190.0000
Mean :	80.9401
Stand. Deviation :	20.5185
df :	230399

Original histogram of DN values for Howe Hill TM band 4



Histogram of
CONTIR

Class width :	1.0000
Display minimum :	0.0000
Display maximum :	255.0000
Actual minimum :	0.0000
Actual maximum :	255.0000
Mean :	108.5659
Stand. Deviation :	27.6579
df :	230399

Histogram of DN values resulting from an applied linear stretch of 5%

Appendix B

Error Matrix Analysis of SUPER5 (columns : truth) against SUPER51 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	458	22	6	0	27	513	0.1072
2	3	154	24	4	54	239	0.3556
3	63	95	658	127	276	1219	0.4602
4	7	16	183	350	1046	1602	0.7815
5	4	39	156	42	2378	2619	0.0920
Total	535	326	1027	523	3781	6192	
ErrorO	0.1439	0.5276	0.3593	0.3308	0.3711		0.3543

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

Error Matrix Analysis of SUPER5 (columns : truth) against SUPER52 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	457	22	7	0	28	514	0.1109
2	5	150	31	4	57	247	0.3927
3	68	104	711	299	473	1655	0.5704
4	0	4	103	170	742	1019	0.8332
5	5	46	175	50	2481	2757	0.1001
Total	535	326	1027	523	3781	6192	
ErrorO	0.1458	0.5399	0.3077	0.6750	0.3438		0.3590

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

Error Matrix Analysis of SUPER5 (columns : truth) against SUPER53 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	460	20	5	0	28	513	0.1033
2	5	157	31	4	57	254	0.3819
3	64	97	713	290	479	1643	0.5660
4	0	4	102	179	746	1031	0.8264
5	6	48	176	50	2471	2751	0.1018
Total	535	326	1027	523	3781	6192	
ErrorO	0.1402	0.5184	0.3057	0.6577	0.3465		0.3572

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

Error Matrix Analysis of COMBINED (columns : truth) against UNSUPER2 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	478	29	5	0	33	545	0.1229
2	46	278	198	26	493	1041	0.7329
3	11	2	406	209	140	768	0.4714
4	0	3	46	157	644	850	0.8153
5	0	14	252	131	2591	2988	0.1329
Total	535	326	907	523	3901	6192	
ErrorO	0.1065	0.1472	0.5524	0.6998	0.3358		0.3685

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

Error Matrix Analysis of COMBINED (columns : truth) against COMBO (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	421	4	1	0	15	441	0.0454
2	3	142	21	5	53	224	0.3661
3	13	3	504	143	160	823	0.3876
4	90	125	257	341	1501	2314	0.8526
5	8	52	124	34	2172	2390	0.0912
Total	535	326	907	523	3901	6192	
ErrorO	0.2131	0.5644	0.4443	0.3480	0.4432		0.4218

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

Error Matrix Analysis of COMBINED (columns : truth) against COMBO2 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	382	2	0	0	12	396	0.0354
2	3	142	21	5	53	224	0.3661
3	13	3	504	143	160	823	0.3876
4	129	127	258	341	1504	2359	0.8554
5	8	52	124	34	2172	2390	0.0912
Total	535	326	907	523	3901	6192	
ErrorO	0.2860	0.5644	0.4443	0.3480	0.4432		0.4281

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

Error Matrix Analysis of COMBINED (columns : truth) against COMBO3 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	382	2	0	0	12	396	0.0354
2	3	142	21	5	53	224	0.3661
3	7	2	472	126	135	742	0.3639
4	135	128	288	357	1527	2435	0.8534
5	8	52	126	35	2174	2395	0.0923
Total	535	326	907	523	3901	6192	
ErrorO	0.2860	0.5644	0.4796	0.3174	0.4427		0.4304

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)